

HALF-REGION DEPTH FOR STOCHASTIC PROCESSES

JAMES KUELBS AND JOEL ZINN

ABSTRACT. We study the concept of half-region depth, introduced by López-Pintado and Romo in [LPR11]. We show that for a wide variety of standard stochastic processes, such as Brownian motion and other symmetric stable processes with stationary independent increments tied down at 0, half-region depth assigns depth zero to all sample functions. To alleviate this difficulty we introduce a method of smoothing, which often not only eliminates the problem of zero depth, but allows us to extend the theoretical results on consistency in that paper up to the \sqrt{n} level for many smoothed processes.

1. INTRODUCTION AND SOME NOTATION

A number of depth functions are available to provide an ordering of finite dimensional data, and more recently in [LPR11] the interesting notion of half-region depth for stochastic processes was introduced. This depth applies to data given in terms of infinite sequences, as functions defined on some interval, and even in more general settings. However, as we will see, one must exercise some care in its use.

In this paper we focus on three issues. The first is to show (see section 2) that for many standard data sources this depth is identically zero, and hence the need to be cautious when employing it. A second issue we examine is how the problem of zero half-region depth can be avoided, and in Proposition 4 it is shown that a smoothing of the data process will eliminate this problem. The third issue we consider involves limit theorems for the empirical half-region depth of these smoothed processes, and Theorem 1 is a basic consistency result with Theorem 2 and Corollary 6 providing some rates of convergence for this consistency. In fact, it provides a sub-Gaussian tail bound. Now we turn to the notation used throughout the paper. In the remainder of this section we indicate some additional details as to how these issues are addressed, and how our results relate to other recent papers.

To fix some notation let $X := \{X(t) = X_t : t \in T\}$ be a stochastic process on the probability space (Ω, \mathcal{F}, P) , all of whose sample paths are in $M(T)$, a linear space of real valued functions on T which we assume to contain the constant functions. To handle measurability issues, we also **always** assume

2010 *Mathematics Subject Classification.* Primary 60F05; Secondary 60F17, 62E20.

Key words and phrases. depth, consistency, central limit theorems, empirical processes.

*Partially supported by NSF grant DMS-1208962.

that $h \in M(T)$ implies

$$(1) \quad \sup_{t \in T} h(t) = \sup_{t \in T_0} h(t) < \infty,$$

where T_0 is a fixed countable subset of T . Typical examples of $M(T)$ are the uniformly bounded continuous functions on T when T is a separable metric space, or the space of cadlag functions on T for T a compact interval of the real line. In either of these situations T_0 could be any countable dense subset of T . It should also be observed that since (1) holds on the linear space $M(T)$, then $h \in M(T)$ implies

$$(2) \quad \inf_{t \in T} h(t) = \inf_{t \in T_0} h(t) > -\infty \text{ and } \|h\|_\infty \equiv \sup_{t \in T} |h(t)| = \sup_{t \in T_0} |h(t)| < \infty.$$

If $g, h: T \rightarrow \mathbb{R}$ and $S \subseteq T$, let $g \preceq_S h$ (resp., $g \succeq_S h$), denote that $g(t) \leq h(t)$ (resp., $h(t) \geq g(t)$) for all $t \in S$. When $S = T$ we will simply write $g \preceq h$ (resp., $g \succeq h$). Then, for a function $h \in M(T)$, the half-region depth with respect to P is defined as

$$(3) \quad D(h, P) := D_{HR}(h, P) := \min(P(X \succeq h), P(X \preceq h)).$$

To simplify, we also will write $D(h)$ for $D(h, P)$ when the probability measure P is understood. Since $M(T)$ is a linear space with (1) and (2) holding, and the sample paths of the stochastic process X are in $M(T)$, we see for each $h \in M(T)$ that

$$(4) \quad \{X \preceq h\} = \{X \preceq_{T_0} h\} \text{ and } \{X \succeq h\} = \{X \succeq_{T_0} h\}.$$

Thus the events in (3) are in \mathcal{F} and the probabilities are defined.

Let X_1, X_2, \dots be i.i.d. copies of the process X , and assume X, X_1, X_2, \dots are defined on the probability space (Ω, \mathcal{F}, P) suitably enlarged, if necessary, and that all sample paths of each X_j are in $M(T)$. Then, the empirical half-region depth of $h \in M(T)$ based on the i.i.d. copies X_1, \dots, X_n is given by

$$(5) \quad D_n(h) = \min\left\{\frac{1}{n} \sum_{j=1}^n I(X_j \succeq h), \frac{1}{n} \sum_{j=1}^n I(X_j \preceq h)\right\}.$$

It is not surprising that like many other infinite dimensional problems, half-region depth is fraught with difficulties not found in the finite dimensional setting. In the next section we examine one such difficulty, namely, that there are classical situations in which the half-region depth is equal to zero for all $h \in M(T)$. The recent paper [DGC11] obtains a result of similar type for Tukey's half-space depth in the sequence space ℓ_2 , and here we'll present half-region depth examples that include data that appears as random sequences, and also as random functions from familiar continuous time stochastic processes. In particular, we will see sample continuous Brownian motion, tied down to be zero at $t = 0$ with probability one, assigns zero half-region depth to all functions $h \in C[0, 1]$, but this sort of problem also holds for other continuous time processes widely used to model data in a

variety of settings. Hence without suitable care, in these situations one is dealing with an object with little significance in the sense that if X is a stochastic process with sample paths in $M(T)$, and all $h \in M(T)$ have zero half-region depth with respect to $P = \mathcal{L}(X)$, then the implications for empirical consistency and central limit type behavior are trivial. That is, if the half-region depth function of every point $h \in M(T)$ is zero, then given h , either $I(X(t) \geq h(t) \forall t \in [0, 1]) = 0$ a.s. or $I(X(t) \leq h(t) \forall t \in [0, 1]) = 0$ a.s. with respect to P . Since the empirical half-space depth function $D_n(h)$ given in (5) is based on the minimum of two sums of such things, we have $D_n(h) - D(h) = D_n(h) = 0$ a.s. with respect to P . Hence even the CLT is degenerate in this case.

Fortunately, in the final proposition of the next section we will see that in many situations smoothing the process by adding an independent real valued random variable Z with a density as in (21) changes things dramatically for half-region depth. For example, sample continuous Brownian motion then attributes strictly positive depth to continuous functions. In later sections we also present additional positive results for this depth. These include consistency results, and also some asymptotics at the \sqrt{n} level. In these results the process X will be as in (21), and also satisfy some additional assumptions. It may also be worthwhile to mention that perhaps other forms of smoothing would be more suitable for other types of depth. This comment is motivated by the zero Tukey-depth result in [DGC11], and also the zero projection depth results in [CC14], which we found as we were in the final writing of this paper. Hence, it would be of interest to determine if a method of smoothing, of one sort or another, can be found to bypass this difficulty in other situations.

In contrast to the smoothing we use, the paper [LPR11] also presents an alternative called modified half-region depth, which is non-degenerate at zero. There the depth itself is changed so as to be less restrictive, whereas here we retain the depth, but apply it to data which has been smoothed as in (21). Moreover, the zero depth results we obtain are such that every function in the natural support of the process has depth zero, i.e. for sample continuous Brownian motion starting at zero at time zero, every continuous path has half-region depth zero. The results in [DGC11] and [CC14] differ in that they show almost every function has zero depth with respect to the law of the process. Finally, we point out that the size of the collection of evaluation maps used in formulating a depth in the infinite dimensional setting, can make an enormous difference. If the collection is too large it is likely the depth will be degenerate, and if it is too small the depth may not reveal details of importance in the data. This phenomenon also appears in connection with the central limit theorems we obtained for empirical processes and empirical quantile processes in [KKZ13] and [KZ13b], where these CLTs may fail if the class of sets is too large, or there are degeneracies in the sample paths, as with Brownian motion tied down at zero. Again, smoothing helps, but one still needs to be careful, since the exact form of the

depth and the evaluation maps used to define it can still produce unusual behavior. For example, in the setting of half-region depth the symmetric stable processes with stationary independent increments, cadlag paths on $[0, 1]$, and tied down at $t = 0$, are such that all cadlag paths on $[0, 1]$ have half-region depth zero (Corollary 4 below), whereas by Proposition 4 these processes smoothed as in (21) have positive depth. Moreover, they satisfy the consistency results and \sqrt{n} -asymptotics provided in Theorems 1, 2, and 3. However, if we look at the increment half-region depth formed by differences of evaluations over only countably many disjoint subintervals of $[0, 1]$ as in Corollary 3, we see that both the smoothed and the unsmoothed version of these processes yield zero increment half-region depth for every function on $[0, 1]$. Of course, similar comments apply to sample continuous Brownian motion, and we also have the half-region depth as defined in Corollary 2 degenerate at zero for all continuous functions on $[0, 1]$ for both the smoothed and unsmoothed versions of Brownian motion.

2. ZERO HALF-REGION DEPTH AND HOW IT CAN BE ELIMINATED

The gist of this section is that for many stochastic processes used in modeling data, half-region depth may be identically zero, but if we smooth the processes as indicated in Proposition 4, this problem is eliminated.

Subsection 2.1 deals with explicit classes of examples, and although these results demonstrate that zero half-region depth is a common phenomenon for many standard processes, the tools developed there should be useful when examining other processes for this problem. Furthermore, it should also be observed that the smoothing result in subsection 2.2, and the consistency and \sqrt{n} -asymptotics of sections 3 and 4, are independent of the proofs in subsection 2.1.

2.1. Some Examples. The half-region depths we examine first are for product probabilities P on the space of all real sequences $R(T)$, where $T = \{t : t = 1, 2, \dots\}$, and for each $h \in R(T)$ the half-region depth remains to be defined as in (3). As before we will write $D(h)$ for $D(h, P)$ when the probability measure P is understood.

For many such P the uniformly bounded sequences $M(T)$ have probability zero, yet we still want to examine such situations as they are natural models of data sources, and they also can be used (as in Corollaries 2 and 3) to determine when a half-region depth may be zero. For example, if P is the product probability whose coordinates are i.i.d. centered Gaussian with variance one, then every coordinate-wise bounded sequence in $R(T)$ has half-region depth equal to zero with respect to this P . Although the set of all such sequences has P -probability zero in this example, a little thought suggests much more may be true, and our next proposition shows that under rather broad circumstances the half-region depth may be zero for all sequences in $R(T)$. In particular, it applies to the Gaussian example we mentioned, and in Corollary 1 it also allows us to examine the situation for sequences

converging to zero, which are relevant when P assigns mass one to a Banach sequence space such as c_0 or ℓ_p , $1 \leq p < \infty$.

Furthermore, if P assigns probability one to $M(T)$, then using Proposition 4 at the end of this section we can show the half-region depth of every $h \in M(T)$ can be strictly positive for a smoothed version of the input data. This latter result applies to data indexed by countable or uncountable T , and $M(T)$ is as defined earlier. Of course, if T is countably infinite, then $M(T)$ is a subset of the sequence space ℓ_∞ , but our results also apply to many standard stochastic processes indexed by uncountable T .

Our first result provides necessary and sufficient conditions for half-region depth to be identically zero for P a product measure on the sequence space $R(T)$. In contrast, a sufficient condition that implies a half-space depth is zero with P -probability one in $R(T)$ for various probabilities P , can be found in [KZ13a]. However, these half-space depths are not zero everywhere, so determining when they are zero, when they are positive, and consistency issues for the related empirical depth are the main concerns there.

Proposition 1. Let $\{Z_t : t \geq 1\}$ be independent rv's on the probability space (Ω, \mathcal{F}, P) with distribution functions F_t , and assume $\mathbf{a} = \{a_t\}_{t=1}^\infty$ is any sequence in $R(T)$. Then,

$$D(\mathbf{a}, P) = 0$$

if and only if

(i) for at least one $t \in T$, $P(Z_t \geq a_t) = 0$ or $P(Z_t \leq a_t) = 0$, or

(ii) for all $t \in T$, $P(Z_t \geq a_t) > 0$ and $P(Z_t \leq a_t) > 0$, and

$$(6) \quad \sum_{t \in T} P(Z_t \neq a_t) = \infty.$$

Remark 1. Under the conditions of Proposition 1, it is immediate that the the conclusion of Proposition 1 is equivalent to the claim that

$$D(\mathbf{a}, P) > 0$$

if and only if for all $t \in T$, $P(Z_t \geq a_t) > 0$ and $P(Z_t \leq a_t) > 0$, and

$$(7) \quad \sum_{t \in T} P(Z_t \neq a_t) < \infty.$$

Proof. Under the assumptions of Proposition 1, it suffices to prove Remark 1. To do this we first we note that

$$\begin{aligned} D(\mathbf{a}, P) &= \min(P(Z_t \leq a_t \text{ for all } t \geq 1), P(Z_t \geq a_t \text{ for all } t \geq 1)) \\ &= \min\left(\prod_{t \geq 1} F_t(a_t), \prod_{t \geq 1} (1 - F_t^-(a_t))\right), \end{aligned}$$

where $F_t^-(x)$ is the left limit at $x \in \mathbb{R}$.

Hence, $D(\mathbf{a}, P) > 0$ if and only if for all $t \in T$ we have $P(Z_t \geq a_t) > 0$ and $P(Z_t \leq a_t) > 0$, and both the products

$$(8) \quad \prod_{t \geq 1} F_t(a_t) = \prod_{t \geq 1} (1 - P(Z_t > a_t)),$$

and

$$(9) \quad \prod_{t \geq 1} (1 - F_t^-(a_t)) = \prod_{t \geq 1} (1 - P(Z_t < a_t))$$

are strictly positive. Since $P(Z_t \geq a_t) > 0$ and $P(Z_t \leq a_t) > 0$ for all $t \in T$, the products in (8) and (9) are strictly positive if and only if

$$(10) \quad \sum_{t \in T} P(Z_t > a_t) < \infty,$$

and

$$(11) \quad \sum_{t \in T} P(Z_t < a_t) < \infty,$$

respectively. Now (10) and (11) holding is equivalent to (7), and hence the proof is complete. \square

Corollary 1. Let $\{Z_t : t \geq 1\}$ be independent rv's on the probability space (Ω, \mathcal{F}, P) with continuous distribution functions F_t for $t \in T_1$, where T_1 is an infinite subset of T . Then,

$$D(\mathbf{a}, P) = 0$$

for **all** sequences $\mathbf{a} = \{a_t\}_{t=1}^\infty$ in $R(T)$. Furthermore, if for $t \in T_1$ and some $\delta > 0$ we weaken the continuity assumption to F_t being continuous on $(-\delta, \delta)$, then $D(\mathbf{a}, P) = 0$ for **all** sequences $\mathbf{a} = \{a_t\}_{t=1}^\infty$ such that $\lim_{t \rightarrow \infty} |a_t| = 0$.

Proof. If the distribution functions F_t are continuous on \mathbb{R} for all $t \in T_1$, where T_1 is an infinite subset of T , then $P(Z_t \neq a_t) = 1$ for all such t 's and (6) holds. Thus $P(Z_t \geq a_t) > 0$ and $P(Z_t \leq a_t) > 0$ for all $t \in T$, and part (ii) of Proposition 1, implies $D(\mathbf{a}, P) = 0$. Of course, if it is not the case that $P(Z_t \geq a_t) > 0$ and $P(Z_t \leq a_t) > 0$ for all $t \in T$, then we also have $D(\mathbf{a}, P) = 0$.

If the assumption of continuity is weakened as indicated, then an entirely similar argument applies for all sequences converging to zero. \square

Remark 2. In the previous corollary continuity of the distributions $F_t, t \in T$, played an important role in showing zero half-region depth, but it clearly is not a necessary condition. For example, if $\{Z_t : t \in T\}$ are independent random variables with $P(Z_t = \pm c_t) = d_t, t \in T$, where $\{c_t : t \in T\}$ are strictly positive constants, $\sum_{t \in T} d_t = \infty$, and $F_t, t \in T$, is arbitrary otherwise, then Proposition 1 immediately implies for any sequence $\mathbf{a} = \{a_t : t \in T\}$

$$D(\mathbf{a}, P) = 0.$$

It is also easy to formulate two immediate consequences of Corollary 1, where natural sequential half-region depths will always be zero for probabilities which behave well in many instances, and are important in many modeling situations. Since more restrictions in the definition of a half-region depth make it easier for the depth to be zero, it is interesting to observe that in both examples the class of evaluation maps used to define the depths is again countably infinite. In the first we assume P is a centered Gaussian probability measure on a separable Banach space with infinite dimensional support. Then, it is well known that there are many sequences of continuous linear functionals $\mathcal{A} = \{\alpha_t : t \in T\} \subseteq B^*$ that are i.i.d. centered Gaussian random variables with $\int_B \alpha_t^2(x) dP(x) = 1$, and for P -almost all $x \in B$

$$\lim_{n \rightarrow \infty} \|x - \sum_{t=1}^n \alpha_t(x) S\alpha_t\| = 0,$$

where $\|\cdot\|$ is the norm on B , and for each $\alpha \in B^*$, $S\alpha$ is the Bochner integral $\int_B x\alpha(x) dP(x)$. Hence, with P -probability one the sequence $\mathcal{A} = \{\alpha_t : t \in T\}$ determines $x \in B$ in the sense that above series converges to x , and we define the \mathcal{A} -half-region depth of a vector $\mathbf{a} \in B$ to be

$$(12) \quad D_{\mathcal{A}}(\mathbf{a}, P) = \min\{P(\alpha_t(x) \geq \alpha_t(\mathbf{a}) \ \forall t \in T), P(\alpha_t(x) \leq \alpha_t(\mathbf{a}) \ \forall t \in T)\}.$$

Corollary 2. If P is a centered Gaussian measure on a separable Banach space with infinite dimensional support, and $\mathcal{A} = \{\alpha_t : t \in T\} \subseteq B^*$ is as above, then for all $\mathbf{a} \in B$

$$(13) \quad D_{\mathcal{A}}(\mathbf{a}, P) = 0.$$

In the second application of Proposition 1 we let $X = \{X(t) : t \in [0, 1]\}$ be a symmetric non-degenerate stable process with stationary independent increments and cadlag sample paths on $[0, 1]$. If X is tied down at $t = 0$, then Proposition 3 below shows that the half-region depth of every cadlag path on $[0, 1]$ is zero with respect to P , and here we examine what might be considered a natural depth for the increments of these processes. Unfortunately, this depth is also zero for every function on $[0, 1]$.

Corollary 3. Let $\mathcal{I} = \{I_j = [u_j, v_j], j \geq 1\}$ consist of disjoint intervals of $[0, 1]$, and define the increment half-region depth for every function h on $[0, 1]$ with respect to $P = \mathcal{L}(X)$ and \mathcal{I} by

$$D_{\mathcal{I}}(h, P) = \min\{P(X(I_j) \geq h(I_j) \ \forall j \geq 1), P(X(I_j) \leq h(I_j) \ \forall j \geq 1)\},$$

where $f(I_j) = f(v_j) - f(u_j)$ for every function f on $[0, 1]$. Then,

$$(14) \quad D_{\mathcal{I}}(\mathbf{a}, P) = 0.$$

As mentioned above, both Corollaries 2 and 3 are immediate from Corollary 1, and the continuity of the relevant distribution functions.

The next proposition will allow us to obtain several more typical examples of “zero half-region depth”.

Proposition 2. Let $\{X(t) : t \in T\}$ and $\{Y(t) : t \in T\}$ be i.i.d stochastic processes on (Ω, \mathcal{F}, P) , all of whose sample paths are in the linear space of functions $M(T)$. If $h \in M(T)$ and

$$(15) \quad P(X - Y \preceq_S 0) = 0$$

for some subset S of T_0 , then $D(h, P) = 0$.

Proof. If the depth of $h \in M(T)$ is positive, then the product, $P(h \preceq X) \cdot P(X \preceq h)$, is positive. So, since we always are assuming (1), (4) and (15), we then have

$$(16) \quad \begin{aligned} 0 &< P(h \preceq_{T_0} X) \cdot P(X \preceq_{T_0} h) = P(h \preceq_{T_0} X, Y \preceq_{T_0} h) \\ &\leq P(Y \preceq_{T_0} X) \leq P(Y - X \preceq_S 0) = 0. \end{aligned}$$

□

Corollary 4. Let X be an independent increment process with paths in the Skorohod space $D[0, 1]$ such that

- (1) the increments have a continuous distribution, and
- (2) $P(X(0) = 0) = 1$.

If $h \in D[0, 1]$, then $D(h, P) = 0$.

Proof. Let $Z = X - Y$, where X and Y are defined on the probability space (Ω, \mathcal{F}, P) , Y is an independent copy of X , and X and Y have sample paths in $D[0, 1]$. Using Proposition 2, with T_0 the rational numbers in $[0, 1]$ and $S = \{\frac{1}{k} : k = 1, 2, \dots\}$, we only have to check that $P(Z \preceq_S 0) = 0$. We'll assume not. But, by the (right) continuity at $t = 0$ and telescoping terms we have

$$(17) \quad Z\left(\frac{1}{k}\right) = \lim_{r \rightarrow \infty} [Z\left(\frac{1}{k}\right) - Z\left(\frac{1}{r+1}\right)] = \lim_{r \rightarrow \infty} \sum_{j=k}^r \Delta_j(Z) = \sum_{j=k}^{\infty} \Delta_j(Z),$$

where $\Delta_j(Z) = [Z(\frac{1}{j}) - Z(\frac{1}{j+1})]$. Therefore, by our choice of S and (17)

$$\begin{aligned} 0 &< P(Z \preceq_S 0) = P\left(\sum_{j=k}^{\infty} \Delta_j(Z) \leq 0, \forall k \geq 1\right) \\ &\leq P\left(\sum_{j=k}^{\infty} \Delta_j(Z) \leq 0, \text{ eventually in } k\right). \end{aligned}$$

This last event is in the tail σ -field of $\{Z(\frac{1}{j}) - Z(\frac{1}{j+1}) : j \geq 1\}$, so by Kolmogorov's zero-one law and the symmetry of Z , we have

$$P\left(\sum_{j=k}^{\infty} \Delta_j(Z) \leq 0, \text{ eventually in } k\right) = P\left(\sum_{j=k}^{\infty} \Delta_j(Z) \geq 0, \text{ eventually in } k\right),$$

and both probabilities are one. Hence, by (17) we have

$$P\left(Z\left(\frac{1}{k}\right) = 0 \text{ eventually in } k\right) = 1,$$

and therefore $P(Z(1/k) - Z(1/k + 1) = 0 \text{ eventually in } k) = 1$. By the independence of the increments this last statement is equivalent to

$$\sum_{k=1}^{\infty} P(Z(1/k) - Z(1/k + 1) \neq 0) < \infty.$$

Since each term is 1, we have a contradiction. \square

Remark 3. Let $X = \{X(t) : t \in [0, 1]\}$ be a symmetric stable process with parameter $r \in (0, 2]$, and stationary independent increments with paths in $D[0, 1]$. If we also have $P(X(0) = 0) = 1$, then the conclusion of Corollary 4 immediately holds. If $r = 2$ and X is Brownian motion with continuous sample paths, then the result also holds in that setting. However, if X is a Poisson process with parameter $\lambda > 0$, then the first condition of Corollary (4) does not hold. And, if ξ has an exponential distribution with mean λ , then

$$P(X(t) \leq 0 \text{ for all } t \in [0, 1]) = P(\xi > 1) > 0.$$

Therefore, the half-space depth of the 0 function is positive. Of course, the same conclusion is valid for compound Poisson processes starting at zero with probability one.

Corollary 5. Let $X = \{X(t) : t = (t_1, t_2) \in [0, 1] \times [0, 1]\}$ be a centered Brownian sheet with covariance

$$(18) \quad E(X(t_1, t_2)X(s_1, s_2)) = \min\{s_1, t_1\} \min\{s_2, t_2\},$$

and continuous paths on $T = [0, 1] \times [0, 1]$. If h is a continuous function on T and P is the law of X , then the half-region depth $D(h, P) = 0$.

Proof. Let $Z = X - Y$, where X and Y are defined on the probability space (Ω, \mathcal{F}, P) , Y is an independent copy of X , and X and Y have sample paths in $C(T)$. Let T_0 be the subset of T consisting of points with both coordinates rational numbers in $[0, 1]$ and let $S = \{t \in T_0 : t = (t_1, t_1)\}$. Using Proposition 2, we only have to check that $P(Z \preceq_S 0) = 0$. We'll assume not. Then,

$$0 < P(Z(t) \leq 0, \text{ for all } t \in S) = P(B(u) \leq 0 \text{ for all } u \in [0, 1] \cap Q),$$

where Q is the rational numbers and $B(u) = Z(u, u), u \in [0, 1]$. Since $\{B(u) : u \in [0, 1]\}$ is a Brownian motion process with continuous sample paths and $P(B(0) = 0) = 1$, we have

$$P(B(u) \leq 0 \text{ for all } u \in [0, 1] \cap Q) = P(B(u) \preceq_{[0, 1]} 0) = 0,$$

where the last equality follows from Remark 3. \square

The next result applies to many Markov processes with or without independent increments.

Proposition 3. Assume the stochastic process $X = \{X_t : 0 \leq t \leq 1\}$ has sample paths in the Skorohod space $D[0, 1]$ and it satisfies

- (1) the Blumenthal zero-one law at $t = 0$, i.e., for every $A \in \mathcal{F}_0^+ := \cap_{t>0} \mathcal{F}_t$ we have $P(A) = 0$ or 1 , where $\mathcal{F}_t = \cup_{0 \leq s \leq t} \sigma(X_s)$ and $\sigma(X_s)$ is the minimal sigma-field making X_s measurable, and
- (2) for every $t > 0$, $X(t)$ has a continuous distribution function.

Then, the half-region depth $D(h, P) = 0$ for every $h \in D[0, 1]$.

Proof. If $D(h, P) >$, then

$$(19) \quad P(X(\cdot) \geq_{[0,1]} h(\cdot)) > 0$$

and

$$(20) \quad P(X(\cdot) \leq_{[0,1]} h(\cdot)) > 0.$$

For $n \geq 1$, let

$$E_n = \{X(\cdot) \geq_{[0,1/n]} h(\cdot)\},$$

and

$$F_n = \{X(\cdot) \leq_{[0,1/n]} h(\cdot)\}.$$

Then, for every integer k

$$E = \{X(t) \geq h(t) \text{ eventually as } t \downarrow 0\} = \cup_{n \geq k} E_n,$$

and

$$F = \{X(t) \leq h(t) \text{ eventually as } t \downarrow 0\} = \cup_{n \geq k} F_n.$$

This implies $E \in \mathcal{F}_{1/k}, F \in \mathcal{F}_{1/k}$ for all $k \geq 1$, and therefore $E, F \in \mathcal{F}_0^+ = \cap_{k=1}^\infty \mathcal{F}_{1/k}$. Now (19) implies $P(E) > 0$ and (20) implies $P(F) > 0$, so the Blumenthal zero-one law implies $P(E) = P(F) = 1$. Since the events E_n and F_n increase in n , we have that there exists a k_0 such that $n \geq k_0$ implies $P(E_n) > 3/4$ and $P(F_n) > 3/4$. Hence

$$P(E_n \cap F_n) > 1/2 \text{ for all } n \geq k_0.$$

Since $E_n \cap F_n = \{X(t) = h(t) \forall t \in [0, 1/n]\}$, this is a contradiction to the fact that $X(t)$ has a continuous distribution for all $t > 0$. Thus the half-region depth of $h \in D[0, 1]$ must be zero. \square

2.2. Eliminating Half-Region Zero Depth By Smoothing. Although sample continuous Brownian motion, tied down to be zero at $t = 0$ with probability one, assigns zero half-region depth to all functions $h \in C[0, 1]$, by starting the process randomly with a density changes things dramatically. This follows immediately from the next proposition, and hence in order to be assured half-region depth is non-trivial, we use smoothing in the results that follow in subsequent sections. Moreover, the precise assumptions used for smoothing in these later results are also important in other parts of their proofs. The smoothed stochastic process $\{X = \{X(t) : t \in T\}$ will be such that

$$(21) \quad X(t) = Y(t) + Z, t \in T,$$

where Z is a real valued random variable independent of the process $Y = \{Y(t) : t \in T\}$, Z has density $f_Z(\cdot)$ on \mathbb{R} , Y has sample paths in the linear

space $M(T)$, and we are assuming $M(T)$ is such that (1) holds. Of course, then (2) also holds, and since we are assuming $M(T)$ contains the constant functions on T , X also has its sample paths in $M(T)$.

Proposition 4. Let $X(t) = Y(t) + Z, t \in T$, where $Y = \{Y(t) : t \in T\}$ has sample paths in the linear space $M(T)$ satisfying (1) and Z is independent of Y with density f_Z . If $f_Z > 0$ a.s. with respect to Lebesgue measure on \mathbb{R} and $h \in M(T)$, then the half-region depth of h determined by $\{X(t) : t \in T\}$ is strictly positive.

Proof. Let $h \in M(T)$. Then,

$$(22) \quad P(X \succeq h) = \int_{-\infty}^{\infty} P(Y(t) \geq h(t) - u \forall t \in T \mid Z = u) f_Z(u) du.$$

Since (1) holds there exists an constant $c > 0$ such that $P(\|Y\|_{\infty} \leq c) > \frac{1}{2}$ and hence for $u > 2c + \|h\|_{\infty}$ we have

$$(23) \quad P(Y(t) \geq h(t) - u \forall t \in T \mid Z = u) \geq P(Y(t) \geq -2c \forall t \in T) > \frac{1}{2}.$$

Since $f_Z > 0$ a.s., by combining (22) and (23) we have

$$P(X \succeq h) \geq \int_{2c + \|h\|_{\infty}}^{\infty} \frac{1}{2} f_Z(u) du > 0.$$

Similarly, $P(X \preceq h) > 0$ for all $h \in M(T)$, and hence $D(h, P) > 0$ for all $h \in M(T)$. \square

Remark 4. If $P(\|Y\|_{\infty} \leq c) > 0$ for all $c > 0$, then it is easy to see from the proof of the previous proposition that the half-region depth could be strictly positive for some $h \in M(T)$ without the density being strictly positive on all of \mathbb{R} .

Remark 5. Let $T = [a, b], -\infty < a < b < \infty$, and assume $M(T)$ denotes the real-valued cadlag paths on T . If $X = \{X(t) : t \in T\}$ has paths in $M(T)$, and $Z := X(a)$ is independent of $\{Y(t) = X(t) - X(a), t \in T\}$ with density $f_Z > 0$ a.s. with respect to Lebesgue measure on \mathbb{R} , then Proposition 4 implies the half-region depth with respect to $P = \mathcal{L}(X)$ is strictly positive on $M(T)$. Hence, under these conditions no smoothing is required to be certain the depth is strictly positive.

3. CONSISTENCY FOR EMPIRICAL HALF-REGION DEPTH

The consistency result we prove depends on two lemmas, which are also important for the \sqrt{n} asymptotics we obtain in Theorem 2. The proof of consistency is an application of empirical process ideas involving the Blum-Dehardt Theorem and bracketing entropy.

Let $X(t) = Y(t) + Z, t \in T$, where $Y = \{Y(t) : t \in T\}$ has sample paths in the linear space $M(T)$ satisfying (1), and Z is independent of Y with density f_Z . Also, assume X_1, X_2, \dots are i.i.d. copies of the process X with sample paths in $M(T)$ and that X, X_1, X_2, \dots are defined on the probability space

(Ω, \mathcal{S}, P) . Then, with half-region depth and half-region empirical depth defined as in (3) and (5), and since for real numbers a, b, c, d

$$(24) \quad |\min\{a, b\} - \min\{c, d\}| \leq |a - c| + |b - d|,$$

the classical strong law of large numbers implies for each $h \in M(T)$

$$(25) \quad \lim_{n \rightarrow \infty} |D_n(h) - D(h)| = 0$$

with probability one. The theorem below refines (25) to be uniform over $h \in E$, where E is a suitably chosen subset of $M(T)$.

Notation 1. For a function $f : \Omega \rightarrow \bar{\mathbb{R}}$ we use the notation f^* to denote a measurable cover function (see Lemma 1.2.1 [vdVW96]).

Theorem 1. Let $X(t) = Y(t) + Z, t \in T$, where $Y = \{Y(t) : t \in T\}$ has sample paths in the linear space $M(T)$, and Z is independent of Y with density $f_Z(\cdot)$ on \mathbb{R} that is absolutely continuous and its derivative $f'_Z(\cdot)$ is in $L_1(\mathbb{R})$. Also, assume X_1, X_2, \dots are i.i.d. copies of the process X with sample paths in $M(T)$ and that X, X_1, X_2, \dots are defined on the probability space (Ω, \mathcal{F}, P) . If E is subset of $M(T)$ such that for every $r > 0$

$$(26) \quad E_r = E \cap \{f \in M(T) : \|f\|_\infty \leq r\}$$

is a sup-norm compact subset of $M(T)$, then with probability one

$$(27) \quad \lim_{n \rightarrow \infty} \sup_{h \in E} |D_n(h) - D(h)|^* = 0.$$

In order to prove this result we first establish some lemmas which will also be useful in our refinements of (27) that follow below. For $h \in M(T)$ we define the stochastic process $\{W_h : h \in M(T)\}$ on (Ω, \mathcal{F}, P) , where

$$(28) \quad W_h \equiv W(h) = \inf_{t \in T} (X(t) - h(t)), h \in M(T).$$

Lemma 1. Let f be a probability density on \mathbb{R} which is absolutely continuous and such that its derivative f' is in $L^1(\mathbb{R})$. Then,

$$(29) \quad \int_{\mathbb{R}} |f(x + \delta) - f(x)| dx \leq |\delta| \int_{\mathbb{R}} |f'(x)| dx.$$

Proof. If $\delta \geq 0$, then

$$\begin{aligned} \int_{\mathbb{R}} |f(x + \delta) - f(x)| dx &= \int_{\mathbb{R}} \left| \int_x^{x+\delta} f'(u) du \right| dx \\ &\leq \int_{\mathbb{R}} \int_x^{x+\delta} |f'(u)| du dx = \int_{\mathbb{R}} \int_{\mathbb{R}} |f'(u)| I_{x \leq u \leq x+\delta} du dx \\ &= (\text{by Fubini}) \int_{\mathbb{R}} |f'(u)| \int_{\mathbb{R}} I_{x \leq u \leq x+\delta} dx du = \delta \int_{\mathbb{R}} |f'(u)| du, \end{aligned}$$

which gives (29)). The case $\delta < 0$ follows similarly. \square

Lemma 2. Let X be as in (21) with Y and Z satisfying the assumptions of Theorem 1, and assume W_h be as in (28). Then, for $h_1, h_2 \in M(T)$ we have

$$(30) \quad |W_{h_1} - W_{h_2}| \leq \|h_1 - h_2\|_\infty.$$

Hence, if $\|h_1 - h_2\|_\infty \leq \delta$, then

$$(31) \quad |P(W_{h_1} \geq x) - P(W_{h_2} \geq x)| \leq P(x - \delta \leq W_{h_1} \leq x) + P(x - \delta \leq W_{h_2} \leq x),$$

and we also have

$$(32) \quad |P(W_{h_1} \geq x) - P(W_{h_2} \geq x)| \leq 2\delta \int_{\mathbb{R}} |f'_Z(x)| dx.$$

Proof. First observe that

$$\begin{aligned} \inf_{t \in T} (X(t) - h_1(t)) &\leq X(s) - h_1(s) = X(s) - h_2(s) + h_2(s) - h_1(s) \\ &\leq X(s) - h_2(s) + \|h_2 - h_1\|_\infty \text{ for all } s \in T. \end{aligned}$$

Hence,

$$\begin{aligned} W_{h_1} &= \inf_{t \in T} (X(t) - h_1(t)) \leq \inf_{s \in T} (X(s) - h_2(s)) + \|h_2 - h_1\|_\infty \\ &= W_{h_2} + \|h_2 - h_1\|_\infty. \end{aligned}$$

and interchanging h_1 and h_2 we have (30).

Hence, if $\|h_1 - h_2\|_\infty \leq \delta$, we then have from (30) that

$$(33) \quad P(W_{h_1} \geq x) \leq P(W_{h_2} \geq x) + P(x - \delta \leq W_{h_2} \leq x)$$

and

$$(34) \quad P(W_{h_2} \geq x) \leq P(W_{h_1} \geq x) + P(x - \delta \leq W_{h_1} \leq x),$$

and (33) and (34) combine to give (31).

To verify (32) we define for $h \in M(T)$

$$(35) \quad F(h, x) = P(W_h \geq x).$$

From (21), $F(h, x) = P(\inf_{t \in T} (Y(t) - h(t)) + Z \geq x)$, and hence the independence of Y and Z implies

$$(36) \quad F(h, x) = \int_{\mathbb{R}} P_Y(\inf_{t \in T} (Y(t) - h(t)) \geq x - y) f_Z(y) dy.$$

Letting $\xi(h) = \inf_{t \in T} (Y(t) - h(t))$, we see (36) implies

$$\begin{aligned} (37) \quad F(h, x_1) - F(h, x_2) &= \int_{\mathbb{R}} [P_Y(\xi(h) \geq x_1 - y) - P_Y(\xi(h) \geq x_2 - y)] f_Z(y) dy \\ &= \int_{\mathbb{R}} [P_Y(\xi(h) \geq s) [f_Z(x_1 - s) - f_Z(x_2 - s)]] ds. \end{aligned}$$

Therefore,

$$(38) \quad |F(h, x_1) - F(h, x_2)| \leq \int_{\mathbb{R}} |f_Z(x_1 - s) - f_Z(x_2 - s)| ds,$$

and setting $u = x_1 - s$ we have $x_2 - s = (x_2 - x_1) + u$, so Lemma 1 implies

$$|F(h, x_1) - F(h, x_2)| \leq \int_{\mathbb{R}} |f_Z(u) - f_Z(u + (x_2 - x_1))| du \leq |x_1 - x_2| \int_{\mathbb{R}} |f'_Z(x)| dx.$$

Thus the lemma is proven since (31) and the above combine to give (32) when $\|h_1 - h_2\|_{\infty} \leq \delta$. \square

Proof. In order to verify (27) we first will show for every $\epsilon > 0$ there is an $r_0 < \infty$ such that the strong law of large numbers and implies with probability one that

$$(39) \quad \limsup_{n \rightarrow \infty} \sup_{\{h: \|h\|_{\infty} \geq r_0\}} D_n(h) \leq P(\|X\|_{\infty} \geq r_0) \leq \epsilon,$$

and

$$(40) \quad \lim_{r \rightarrow \infty} \sup_{\{h: \|h\|_{\infty} \geq r\}} D(h) = 0.$$

The argument for (39) and (40) is essentially the proof of Proposition 5 in [LPR11], but the details are included below.

To prove (40) we observe

$$\sup_{\|h\|_{\infty} \geq r} D(h) \leq A_r + B_r,$$

where

$$A_r = \sup_{\|h\|_{\infty} \geq r, \|h\|_{\infty} = \sup_{t \in T} h(t)} P(X \succeq h),$$

and

$$B_r = \sup_{\|h\|_{\infty} \geq r, \|h\|_{\infty} = \sup_{t \in T} (-h(t))} P(X \preceq h).$$

Thus

$$\begin{aligned} A_r &\leq \sup_{\|h\|_{\infty} = \sup_{t \in T} h(t) \geq r} P(\sup_{t \in T} X(t) \geq \sup_{t \in T} h(t)) \\ &\leq \sup_{\|h\|_{\infty} = \sup_{t \in T} h(t) \geq r} P(\|X\|_{\infty} \geq \|h\|_{\infty}) \leq P(\|X\|_{\infty} \geq r), \end{aligned}$$

and

$$\begin{aligned} B_r &\leq \sup_{\|h\|_{\infty} = \sup_{t \in T} (-h(t)) \geq r} P(\inf_{t \in T} X(t) \leq \inf_{t \in T} h(t)) \\ &\leq \sup_{\|h\|_{\infty} = \sup_{t \in T} (-h(t)) \geq r} P(\|X\|_{\infty} \geq \|h\|_{\infty}) \leq P(\|X\|_{\infty} \geq r), \end{aligned}$$

and hence we have (40). To prove (39) we note that

$$\sup_{\|h\|_{\infty} \geq r} D_n(h) \leq A_{r,n} + B_{r,n}.$$

where

$$A_{r,n} = \sup_{\|h\|_\infty \geq r, \|h\|_\infty = \sup_{t \in T} h(t)} \frac{1}{n} \sum_{j=1}^n I(X_j \succeq h),$$

and

$$B_{r,n} = \sup_{\|h\|_\infty \geq r, \|h\|_\infty = \sup_{t \in T} (-h(t))} \frac{1}{n} \sum_{j=1}^n I(X_j \preceq h).$$

Thus, in similar fashion it follows that

$$A_{r,n} \leq \sup_{\|h\|_\infty = \sup_{t \in T} h(t) \geq r} \frac{1}{n} \sum_{j=1}^n I(\|X_j\|_\infty \geq \|h\|_\infty) \leq \frac{1}{n} \sum_{j=1}^n I(\|X_j\|_\infty \geq r),$$

and

$$B_{r,n} \leq \sup_{\|h\|_\infty = \sup_{t \in T} (-h(t)) \geq r} \frac{1}{n} \sum_{j=1}^n I(\|X_j\|_\infty \geq \|h\|_\infty) \leq \frac{1}{n} \sum_{j=1}^n I(\|X_j\|_\infty \geq r),$$

and therefore we have (39).

Since $\epsilon > 0$ is arbitrary, (40) and (39) combine to imply (27) provided we show for every $r > 0$ that with probability one

$$(41) \quad \lim_{n \rightarrow \infty} \sup_{h \in E_r} |D_n(h) - D(h)|^* = 0,$$

where E_r is defined as in (26). The proof of (41) follows from the Blum-Dehardt Theorem using the bracketing entropy for E_r as in [Dud99], p. 235. That is, since E_r is compact in $M(T)$ with respect to the sup-norm, for every $\delta > 0$ implies there exists finitely many points $\{h_1, \dots, h_{k(\delta)}\} \subseteq E_r$ such that

$$\sup_{h \in E_r} \inf_{h_j} \|h - h_j\|_\infty \leq \delta.$$

In addition, the brackets $F(\delta, h_j) = \{z \in M(T) : h_j(t) - \delta \leq z(t) \leq h_j(t) + \delta\}$ have union covering E_r with $z \in F(\delta, h_j)$ implying

$$I(X \succeq h_j + \delta) \leq I(X \succeq z) \leq I(X \succeq h_j - \delta).$$

Hence, for $\epsilon > 0$ fixed, and $\delta = \delta(\epsilon) > 0$ such that $4\delta \int_{\mathbb{R}} |f'_Z(x)| dx \leq \epsilon$, we have from (32) that

$$\mathcal{E}_r \equiv \{I(X \succeq h) : h \in E_r\}$$

is a subset of

$$\bigcup_{j=1}^{k(\delta(\epsilon))} \{I(X \succeq z) : z \in E_r, I(X \succeq h_j + \delta) \leq I(X \succeq z) \leq I(X \succeq h_j - \delta)\},$$

and

$$\|I(X \succeq h_j - \delta) - I(X \succeq h_j + \delta)\|_1 \leq \epsilon,$$

where $\|\cdot\|_1$ denotes the L_1 norm with respect to P . Hence, for every $\epsilon > 0$ we have \mathcal{E}_r covered by finitely many L_1 -brackets of diameter ϵ . A similar argument can be made for

$$\mathcal{F}_r \equiv \{I(X \preceq h) : h \in E_r\},$$

and hence (41) holds by (24) and the Blum-Dehardt result mentioned above. Combining (40), (39), and (41) we have (27). \square

3.1. Some Remarks on the C1 Condition in [LPR11]. Let $X = \{X_t : t \in [0, 1]\}$ be a sample continuous stochastic process, and assume P is the Borel probability on $C[0, 1]$ induced by X . The main focus of the paper [LPR11] is the formulation of a consistency result for half-region depth that is uniform over an equicontinuous family of functions on $[0, 1]$, where the depth is with respect to X , or equivalently the probability distribution P . One of the crucial assumptions in this endeavor is that P satisfy their C1 condition, where

C1: Given $\epsilon > 0$, there exists a $\delta > 0$, such that for every pair of functions $h_1, h_2 \in C[0, 1]$ with $\|h_1 - h_2\|_\infty \leq \delta$ implies

$$(42) \quad P(h_1 \preceq_{[0,1]} X \preceq_{[0,1]} h_2) < \epsilon.$$

This condition appears on the bottom of page 1687 in [LPR11]. The notation in [LPR11] is slightly different than that above, but (42) is consistent with their use of C1 on page 1688 of [LPR11]. However, the main problem with (42) as used in [LPR11] is two-fold. First, in their proof of Theorem 3 of [LPR11] it is applied to functions h_1, h_2 which are not continuous, and secondly it is claimed that for $h_1, h_2 \in C[0, 1]$ with $h_1 \preceq h_2$

$$(43) \quad P(h_1 \preceq X) - P(h_2 \preceq X) = P(h_1 \preceq X \preceq h_2),$$

which is far from being true since

$$(44) \quad P(h_1 \preceq X) - P(h_2 \preceq X) = P(\{h_1 \preceq X\} \cap \{h_2 \preceq X\}^c).$$

Hence there are some major concerns with their proof, and in Theorem 1 above we obtained a result that alleviates such concerns. Moreover, we have taken care to discuss when half-region is non-trivial, and how to eliminate the problem of it being trivial by using smoothing. Another question one might ask is whether the quantity

$$(45) \quad |P(h_1 \preceq X) - P(h_2 \preceq X)|,$$

can be made arbitrarily small when $h_1, h_2 \in C[0, 1]$ provided $\|h_1 - h_2\|_\infty$ is sufficiently small. This is an important ingredient in our proof, and we established sufficient conditions for the continuity posed for (45) in Lemma (2), but it is easy to see it may fail in some cases. We conclude this section with two such examples, and for ease in writing we will refer to the continuity posed for (45) as condition C2.

The first example where condition C2 fails is for

$$X_t = \max\{0, B_t\}, t \in [0, 1],$$

where $\{B(t) : t \in [0, 1]\}$ is a sample continuous Brownian motion such that $P(B(0) = 0) = 1$. Thus for $h_1(t) = 0, t \in [0, 1]$, and $h_2(t) = \delta > 0, t \in [0, 1]$, we have

$$P(h_1 \preceq X) - P(h_2 \preceq X) = 1,$$

no matter how small the constant δ . If we let

$$X_t = Z + \max\{0, B_t\},$$

where Z is independent of the Brownian motion B , then for $h_1(t) = c, t \in [0, 1]$, and $h_2(t) = c + \delta > 0, t \in [0, 1]$, an easy calculation implies

$$P(h_1 \preceq X) - P(h_2 \preceq X) = P(c \leq Z < c + \delta).$$

Hence, if $P(Z = c) > 0$, then again no matter how small the constant δ , the condition C2 fails. If Z has a continuous distribution, then C2 holds for all choices of $h_1, h_2 \in C[0, 1]$ constant functions, but that it actually does or does not satisfy C2 is not at all obvious.

The second example fails C2 for the same reasons as those in the previous one, i.e. wherever the process starts, it never goes below that level, and if it starts at a fixed point which has positive probability, then C2 fails. However, it differs in that its paths are Lip-1 with probability one. The example is

$$X_t = \int_0^t N(s) ds, t \in [0, 1],$$

where $N = \{N(t) : t \in [0, 1]\}$ is a Poisson process with parameter one and cadlag paths. Properties such as those mentioned for example one also hold here. The details are left to the reader.

4. ADDITIONAL ASYMPTOTICS FOR HALF-REGION DEPTH

Our next result shows the consistency result of (27) can be refined to include rates of convergence provided we restrict the set E to be a sup-norm compact subset of $M(T)$ satisfying the entropy condition

$$(46) \quad \int_{0+} (\log N(E, \epsilon, \|\cdot\|_\infty))^{\frac{1}{2}} \epsilon^{-\frac{1}{2}} d\epsilon < \infty,$$

where $N(E, \epsilon, \|\cdot\|_\infty)$ is the covering number of E with ϵ -balls in the $\|\cdot\|_\infty$ -norm. In particular, since the processes

$$(47) \quad \{\sqrt{n}(D_n(h) - D(h)) : h \in E\}, n \geq 1,$$

live in $\ell_\infty(E)$, we examine their asymptotic behavior in that setting, and in Corollary 6 produce sub-Gaussian tail bounds that are uniform in n . The basic notation is as in section 3, and we freely use the empirical process ideas for weak convergence in the space $\ell_\infty(E)$ as presented in [Dud99] and [vdVW96].

In the proof of these results we have need for the stochastic processes $\{H_{n,1,h} : h \in E\}, n \geq 1$, and $\{H_{n,2,h} : h \in E\}, n \geq 1$, where

$$(48) \quad H_{n,1,h} \equiv \frac{1}{\sqrt{n}} \sum_{j=1}^n [I(X_j \succeq h) - P(X_j \succeq h)],$$

and

$$(49) \quad H_{n,2,h} \equiv \frac{1}{\sqrt{n}} \sum_{j=1}^n [I(X_j \preceq h) - P(X_j \preceq h)].$$

The first step of our proof will be to show that each of these processes satisfies the CLT in $\ell_\infty(E)$ with limits that are centered, sample path bounded, Gaussian processes $G_1 = \{G_{1,h} : h \in E\}$ and $G_2 = \{G_{2,h} : h \in E\}$, respectively, that are uniformly continuous on E with respect to their L_2 -distances, and have covariance functions

$$(50) \quad E(G_{1,h_1} G_{1,h_2}) = P(X \succeq h_1, X \succeq h_2) - P(X \succeq h_1)P(X \succeq h_2), h_1, h_2 \in E,$$

and

$$(51) \quad E(G_{2,h_1} G_{2,h_2}) = P(X \preceq h_1, X \preceq h_2) - P(X \preceq h_1)P(X \preceq h_2), h_1, h_2 \in E.$$

In the following theorem these Gaussian processes also appear in connection with the limiting finite dimensional distributions of the centered empirical half-region depth processes given in (47), see (53-55).

Theorem 2. Let $X(t) = Y(t) + Z, t \in T$, where $Y = \{Y(t) : t \in T\}$ has sample paths in the linear space $M(T)$ and Z is independent of Y with density $f_Z(\cdot)$ on \mathbb{R} that is absolutely continuous and its derivative $f'_Z(\cdot)$ is in $L_1(\mathbb{R})$. Also, assume X_1, X_2, \dots are i.i.d. copies of the process X with sample paths in $M(T)$ and that X, X_1, X_2, \dots are defined on the probability space (Ω, \mathcal{F}, P) . If E is a sup-norm compact subset of $M(T)$ satisfying the entropy condition (46), then

$$(52) \quad \lim_{r \rightarrow \infty} \sup_{n \geq 1} P^*(\sup_{h \in E} \sqrt{n} |D_n(h) - D(h)| \geq r) = 0,$$

where P^* denotes the outer probability for subsets of (Ω, \mathcal{F}, P) . Furthermore, there is a stochastic process $\{\Gamma_h : h \in E\}$ such that the finite dimensional distributions of the processes $\{\sqrt{n}(D_n(h) - D(h)) : h \in E\}, n \geq 1$ converge weakly to $\{\Gamma_h : h \in E\}$, where

$$(53) \quad \mathcal{L}(\Gamma_h) = \mathcal{L}(G_{1,h}) \text{ for } h \in E \text{ and } P(X \succeq h) < P(X \preceq h),$$

$$(54) \quad \mathcal{L}(\Gamma_h) = \mathcal{L}(G_{2,h}) \text{ for } h \in E \text{ and } P(X \preceq h) < P(X \succeq h),$$

and

$$(55) \quad \mathcal{L}(\Gamma_h) = \mathcal{L}(\min\{G_{1,h}, G_{2,h}\}) \text{ for } h \in E \text{ and } P(X \succeq h) = P(X \preceq h).$$

Proof. Since (24) holds and X, X_1, X_2, \dots are i.i.d. we have

$$(56) \quad \sqrt{n}|D_n(h) - D(h)| \leq |H_{n,1,h}| + |H_{n,2,h}|.$$

Hence (52) will hold provided we show

$$(57) \quad \lim_{r \rightarrow \infty} \sup_{n \geq 1} P^*(\sup_{h \in E} |H_{n,1,h}| \geq r) = 0,$$

and

$$(58) \quad \lim_{r \rightarrow \infty} \sup_{n \geq 1} P^*(\sup_{h \in E} |H_{n,2,h}| \geq r) = 0.$$

To verify (57) and (58) it suffices to show that the stochastic processes $\{H_{n,1,h} : h \in E\}$ and $\{H_{n,2,h} : h \in E\}$ converge weakly in $\ell_\infty(E)$ to the centered Gaussian processes G_1 and G_2 , respectively. That is, once these CLT's hold, then item (iii) of Theorem 1.3.4 of [vdVW96] provides the conclusion we need.

In order to formulate these CLTs in $\ell_\infty(E)$ we let \mathcal{C} be a family of subsets of $M(T)$ indexed by E , where

$$(59) \quad \mathcal{C} = \mathcal{C}_{\text{inf}} \cup \mathcal{C}_{\text{sup}},$$

$$(60) \quad \mathcal{C}_{\text{inf}} = \{C_h : h \in E\} \text{ and } \mathcal{C}_{\text{sup}} = \{\hat{C}_h : h \in E\},$$

$$(61) \quad C_h = \{z \in M(T) : \inf_{t \in T} (z(t) - h(t)) \geq 0\},$$

and

$$(62) \quad \hat{C}_h = \{z \in M(T) : \sup_{t \in T} (z(t) - h(t)) \leq 0\},$$

Of course, since we are assuming $M(T)$ is a linear space such that (1) holds we have the inf and sup defining the sets C_h and \hat{C}_h , respectively, are the same when $t \in T$ is replaced by $t \in T_0$.

Since $\ell_\infty(E)$ is a separable Banach space only when E is finite, we need to use weak convergence in the non-separable setting, and proceed to verify that \mathcal{C}_{inf} and \mathcal{C}_{sup} are both P -Donsker classes of sets. Then, since a finite union of P -Donsker classes is P -Donsker, we will have \mathcal{C} also P -Donsker.

To show \mathcal{C}_{inf} is P -Donsker we recall the stochastic process $\{W_h : h \in M(T)\}$ on (Ω, \mathcal{F}, P) given in (28). Then, the path $X(t, \cdot)$ is in C_h if and only if $W_h(\cdot) \geq 0$, and we also have $X \succeq h$ on T if and only if $W_h \geq 0$. Therefore, \mathcal{C}_{inf} P -Donsker will imply that the empirical processes $\{H_{n,1,h} : h \in E\}$ as given in (48) converge in distribution on $\ell_\infty(E)$ to a centered Gaussian measure γ_{inf} with separable support in $\ell_\infty(E)$. Furthermore, γ_{inf} is induced by the Gaussian process G_1 as indicated above.

Since (46) holds, for every $\delta > 0$ there are $N_\delta \equiv N(E, \delta, \|\cdot\|_\infty)$ functions h_1, \dots, h_{N_δ} in E such that the brackets

$$(63) \quad F(\delta, h_j) = \{z \in M(T) : h_j(t) - \delta \leq z(t) \leq h_j(t) + \delta\}, j = 1, \dots, N_\delta,$$

have union covering E . Furthermore, $z \in F(\delta, h_j)$ implies

$$I(X \succeq h_j + \delta) \leq I(X \succeq z) \leq I(X \succeq h_j - \delta).$$

Hence, for $\delta > 0$ fixed we have from (32) that

$$\mathcal{E} \equiv \{I(X \succeq h) : h \in E\}$$

is a subset of

$$\cup_{j=1}^{N_\delta} \{I(X \succeq z) : z \in E, I(X \succeq h_j + \delta) \leq I(X \succeq z) \leq I(X \succeq h_j - \delta)\}.$$

Furthermore,

$$\begin{aligned} \|I(X \succeq h_j - \delta) - I(X \succeq h_j + \delta)\|_2^2 &= \|I(X \succeq h_j - \delta) - I(X \succeq h_j + \delta)\|_1 \\ &= P(X \succeq h_j - \delta) - P(X_j \succeq h_j + \delta) \leq 4\delta \int_{\mathbb{R}} |f'_Z(x)| dx, \end{aligned}$$

where $\|\cdot\|_p$ denotes the L_p norm with respect to P , and the inequality follows from (32) and (63). Now

$$\begin{aligned} \int_{0+} (\log N(\mathcal{E}, x, \|\cdot\|_2))^{\frac{1}{2}} dx &= \int_{0+} (\log N(\mathcal{E}, x^2, \|\cdot\|_1))^{\frac{1}{2}} dx \\ &\leq \int_{0+} (\log N(\mathcal{E}, x^2, \|\cdot\|_\infty))^{\frac{1}{2}} dx \end{aligned}$$

where the inequality follows since $\|\cdot\|_1 \leq \|\cdot\|_\infty$ on $M(T)$. Letting $s = x^2$ in the right most integral above and applying (46) we have

$$(64) \quad \int_{0+} (\log N(\mathcal{E}, x, \|\cdot\|_2))^{\frac{1}{2}} dx \leq \int_{0+} (\log N(\mathcal{E}, s, \|\cdot\|_\infty))^{\frac{1}{2}} s^{-\frac{1}{2}} ds < \infty.$$

Hence by Ossiander's CLT with bracketing [Oss87], or as in [Dud99], p 239, we have \mathcal{E} a P -Donsker class of functions, which implies \mathcal{C}_{inf} is a P -Donsker class of sets. Hence the empirical processes $\{H_{n,1,h} : h \in E\}$ given in (48) converge weakly in $\ell_\infty(E)$ to the centered Gaussian process G_1 induced by the Radon Gaussian measure γ_{inf} and has covariance as indicated in (4.5). A similar result holds for the empirical processes $\{H_{n,2,h} : h \in E\}$ given in (49), which therefore satisfy the CLT in $\ell_\infty(E)$ with centered Gaussian limit G_2 . Hence (52) is proven.

The next step of our proof is to show the finite dimensional distributions of the stochastic processes in (47) converge. To check this we set

$$F_n(h) = \frac{1}{n} \sum_{j=1}^n I(X_j \preceq h), \quad F(h) = P(X \preceq h),$$

and

$$G_n(h) = \frac{1}{n} \sum_{j=1}^n I(X_j \succeq h), \quad G(h) = P(X \succeq h).$$

Hence, let $I = I_1 \cup I_2 \cup I_3$, where I_1, I_2, I_3 are disjoint,

$$I_1 = \{h_1, \dots, h_{r_1}\}, \quad I_2 = \{h_{r_1+1}, \dots, h_{r_2}\}, \quad I_3 = \{h_{r_2+1}, \dots, h_r\},$$

and

$$I_1 = \{h \in I : F(h) < G(h)\},$$

$$I_2 = \{h \in I : F(h) > G(h)\},$$

and

$$I_3 = \{h \in I : F(h) = G(h)\}.$$

Setting

$$V_n(h) = \sqrt{n}(D_n(h) - D(h)),$$

we have

$$V_n(h) = \sqrt{n}[\min(F_n(h), G_n(h)) - \min(F(h), G(h))],$$

and since I is an arbitrary subset of E to prove the finite dimensional distributions of the processes in (47) we need to show

$$(V_n(h_1), \dots, V_n(h_r)),$$

converges in distribution on \mathbb{R}^r .

For $n \geq 1$ let

$$U_n(h) = \sqrt{n}(F_n(h) - F(h)), h \in I_1,$$

$$U_n(h) = \sqrt{n}(G_n(h) - G(h)), h \in I_2$$

$$U_n(h) = \sqrt{n} \min(F_n(h) - F(h), G_n(h) - G(h)), h \in I_3,$$

and take

$$N(\omega) = \min\{m \geq 1 : U_n(h_i) = V_n(h_i), 1 \leq i \leq r_2, n \geq m\}.$$

Then, the strong law of large numbers implies $P(N < \infty) = 1$, and $U_n(h) = V_n(h)$ for all $h \in I$ and all $n \geq N$. Therefore,

$$\lim_{n \rightarrow \infty} P(\sup_{h \in I} |U_n(h) - V_n(h)| \geq \epsilon) \leq \lim_{n \rightarrow \infty} P(N > n) = 0,$$

and the convergence of the finite dimensional distributions will hold if we show

$$T_n = u_1 U_n(h_1) + \dots + u_r U_n(h_r)$$

converges in distribution for all $(u_1, \dots, u_r) \in \mathbb{R}^r$. Setting

$$S_n = \sum_{j=1}^{r_1} u_j (F_n(h_j) - F(h_j)) + \sum_{j=r_1+1}^{r_2} u_j (G_n(h_j) - G(h_j)),$$

we have

$$T_n = \sqrt{n} [\min[S_n + u_{r_2+1}(F_n(h_{r_2+1}) - F(h_{r_2+1})), S_n + u_{r_2+1}(G_n(h_{r_2+1}) - G(h_{r_2+1}))]] + \sqrt{n} \sum_{j=r_2+2}^r u_j \min[F_n(h_j) - F(h_j), G_n(h_j) - G(h_j)].$$

If

$$\Lambda(a_1, b_1, a_2, b_2, \dots, a_k, b_k) = \sum_{i=1}^k \min[a_i, b_i],$$

then Λ is continuous from \mathbb{R}^{2k} to \mathbb{R}^k . Therefore, if $k = r - r_2$ with

$$R_n = (a_{n,1}, b_{n,1}, \dots, a_{n,r-r_2}, b_{n,r-r_2})$$

and

$$\begin{aligned} a_{n,1} &= \sqrt{n}(S_n + u_{r_2+1}[F_n(h_{r_2+1}) - F(h_{r_2+1})]), \\ b_{n,1} &= \sqrt{n}(S_n + u_{r_2+1}[G_n(h_{r_2+1}) - G(h_{r_2+1})]), \\ a_{n,i} &= \sqrt{n}u_{r_2+i}[F_n(h_{r_2+i}) - F(h_{r_2+i})], i = 2, \dots, r - r_2, \\ b_{n,i} &= \sqrt{n}u_{r_2+i}[G_n(h_{r_2+i}) - G(h_{r_2+i})], i = 2, \dots, r - r_2, \end{aligned}$$

we have R_n converging weakly to a centered Gaussian random variable, i.e. it is a sum of independent vectors in $\mathbb{R}^{2(r-r_2)}$ whose summands are indicator functions multiplied by u_j 's. Now $\Lambda(R_n) = T_n$, and thus the continuous mapping theorem implies T_n converges in distribution. Since the vector $(u_1, \dots, u_r) \in \mathbb{R}^r$ is arbitrary, the finite dimensional distributions converge. Of course, the claims in (53), (54), and (55) involving the one dimensional distributions are also now proven. \square

Next we turn to a corollary of Theorem 2, which provides sub-Gaussian tail bounds for the convergence to zero in (52). To avoid measurability issues arising in its proof, we assume the set E is countable. Of course, under the assumption (1) and that $M(T)$ is a linear space, we have that the random vectors (stochastic processes)

$$\{D_n(h) - D(h) : h \in E\} \text{ and } H_{n,i} := \{H_{n,i,h} : h \in E\}, i = 1, 2,$$

given in (48) and (49), take values in the Banach space $\ell_\infty(E)$ with norm $\|x\|_\infty = \sup_{h \in E} |x_h|$ for $x = \{x_h\} \in \ell_\infty(E)$. Hence, the assumption E is countable implies these random vectors on (Ω, \mathcal{F}, P) are $\ell_\infty(E)$ valued in the sense used in [LT91], so for the convenience of the reader we freely quote from this single source a number of results used in the proof. However, from a historical point of view it should be observed that an important first step in these results involves the Hoffmann-Jørgensen inequalities obtained in [HJ74], and for series and a.s. normalized partial sums of sequences of independent random vectors, some results of a similar nature appeared in [JM75] and [Kue78].

Notation 2. Let X take values in a Banach space B with norm $\|x\| = \sup_{f \in D} |f(x)|$, $x \in B$, where D is a countable subset of the unit ball of the dual space of B and $f(X)$ is measurable for each $f \in D$. Then, we are in the setting used in Chapter 6 of [LT91], and the ψ_2 -Orlicz norm of $\|X\|$ is given by

$$\|X\|_{\psi_2} = \inf\{c > 0 : E(\exp\{(\frac{\|X\|}{c})^2\}) \leq 2\}.$$

Corollary 6. Let $H_{n,i} := \{H_{n,i,h} : h \in E\}$, $i = 1, 2$, be the stochastic processes in (48) and (49), and for $i = 1, 2$

$$(65) \quad \|H_{n,i}\|_\infty = \sup_{h \in E} |H_{n,i,h}|.$$

Then, under the assumptions of Theorem 2 and that E is countable, we have

$$(66) \quad \hat{k} = \sup_{n \geq 1, i=1,2} E(\|H_{n,i}\|_\infty) < \infty,$$

and there exists an absolute constant $k_2 < \infty$ such that for any $r > 0$

$$(67) \quad \sup_{n \geq 1} P(\sup_{h \in E} \sqrt{n} |D_n(h) - D(h)| \geq r) \leq 4 \exp\{-\alpha r^2\}$$

provided $\alpha > 0$ is sufficiently small that

$$(68) \quad \sqrt{4\alpha} k_2 (\hat{k} + 2) < 1.$$

Proof. From (56)

$$P(\sup_{h \in E} \sqrt{n} |D_n(h) - D(h)| \geq r) \leq P(\|H_{n,1}\|_\infty \geq \frac{r}{2}) + P(\|H_{n,2}\|_\infty \geq \frac{r}{2}).$$

Hence, Markov's inequality implies

$$P(\sup_{h \in E} \sqrt{n} |D_n(h) - D(h)| \geq r) \leq \exp\{-\alpha r^2\} \sum_{i=1}^2 E(\exp\{4\alpha \|H_{n,i}\|_\infty^2\}),$$

and (67) holds provided $\alpha > 0$ is sufficiently small that

$$(69) \quad \sqrt{4\alpha} \sup_{n \geq 1, i=1,2} \|H_{n,i}\|_{\infty, \psi_2} < 1,$$

where we write $\|H_{n,i}\|_{\infty, \psi_2}$ to denote the ψ_2 -norm of $\|H_{n,i}\|_\infty$. Now Theorem 6.21 of [LT91] implies there exists an absolute constant $k_2 < \infty$ such that

$$(70) \quad \|H_{n,i}\|_{\infty, \psi_2} \leq k_2 [E(\|H_{n,i}\|_\infty) + (\sum_{j=1}^n \|\frac{Y_j}{\sqrt{n}}\|_{\infty, \psi_2}^2)^{\frac{1}{2}}],$$

where $\{Y_j : j \geq 1\}$ are independent, mean zero, $\ell_\infty(E)$ valued random vectors with $Y_j = \{I(X_j \succeq h) - P(X_j \succeq h) : h \in E\}$ for $j \geq 1$ when $i = 1$, and $Y_j = \{I(X_j \preceq h) - P(X_j \preceq h) : h \in E\}$ for $j \geq 1$ when $i = 2$. Since $\|Y_j\|_\infty \leq 1$, we have $\|\frac{Y_j}{\sqrt{n}}\|_{\infty, \psi_2} \leq \frac{1}{\sqrt{n \log(2)}}$, and (70) implies

$$(71) \quad \|H_{n,i}\|_{\infty, \psi_2} \leq k_2 [E(\|H_{n,i}\|_\infty) + (\frac{1}{\log(2)})^{\frac{1}{2}}],$$

Now from (69) and (71) we have (67) for $\alpha > 0$ sufficiently small that

$$(72) \quad \sqrt{4\alpha} k_2 [\hat{k} + (\frac{1}{\log(2)})^{\frac{1}{2}}] < 1,$$

provided (66) holds.

Hence, to complete the proof we must prove $\hat{k} < \infty$. To accomplish this we first show

$$(73) \quad \sup_{n \geq 1} E(\|H_{n,1}\|_\infty) < \infty.$$

This follows from Proposition 6.8 of [LT91] applied to the partial sums S_k of the $\{Y_j : j \geq 1\}$ with $p = 1$ provided we show $\sup_{n \geq 1} t_{0,n} < \infty$, where

$$(74) \quad t_{0,n} = \inf\{t > 0 : P(\max_{1 \leq k \leq n} \|\frac{S_k}{\sqrt{n}}\|_\infty > t) \leq \frac{1}{8}\}, n \geq 1.$$

Moreover, E countable implies Ottaviani's inequality is available as in Lemma 6.2 of [LT91], and hence for every $u, v > 0$

$$(75) \quad P(\max_{1 \leq k \leq n} \|\frac{S_k}{\sqrt{n}}\|_\infty > u + v) \leq \frac{P(\|\frac{S_n}{\sqrt{n}}\|_\infty > v)}{1 - \max_{1 \leq k \leq n} P(\|\frac{S_n - S_k}{\sqrt{n}}\|_\infty > u)}.$$

Furthermore, since $\frac{S_n}{\sqrt{n}} = H_{n,1}$ for $n \geq 1$, and the proof of Theorem 2 implies $\{H_{n,1} : n \geq 1\}$ satisfies the central limit theorem in $\ell_\infty(E)$, the Portmanteau Theorem (applied to closed sets) implies there exists $u_0 < \infty$ such that for $u \geq u_0$

$$P(\|\frac{S_m}{\sqrt{m}}\|_\infty \geq u) \leq \frac{1}{2}$$

for all $m \in [m_0, \infty)$. Therefore, there exists $u_1 \in [u_0, \infty)$ such that

$$\sup_{m \geq 1} P(\|\frac{S_m}{\sqrt{m}}\|_\infty \geq u_1) \leq \frac{1}{2},$$

and hence

$$\sup_{n \geq 1} \max_{1 \leq k \leq n} P(\|\frac{S_n - S_k}{\sqrt{n}}\|_\infty \geq u_1) \leq \sup_{n \geq 1} \max_{1 \leq m \leq n} P(\|\frac{S_m}{\sqrt{m}}\|_\infty \geq u_1) \leq \frac{1}{2}.$$

Thus (75) implies for all $v > 0$ and $n \geq 1$ that

$$(76) \quad P(\max_{1 \leq k \leq n} \|\frac{S_k}{\sqrt{n}}\|_\infty > u_1 + v) \leq 2P(\|\frac{S_n}{\sqrt{n}}\|_\infty > v).$$

Again, by the central limit theorem there exists $v_1 < \infty$ such that $v \geq v_1$ implies

$$2 \sup_{n \geq 1} P(\|\frac{S_n}{\sqrt{n}}\|_\infty \geq v) \leq \frac{1}{8},$$

and hence we see from (76) that $\sup_{n \geq 1} t_{0,n} \leq u_1 + v_1 < \infty$ when $i = 1$ (and the partial sums come from the $\{Y_j : j \geq 1\}$). However, the same proof applies when $i = 2$ and the partial sums are formed from $\{Z_j : j \geq 1\}$, where $Z_j = \{I(X_j \preceq h) - P(X_j \preceq h) : h \in E\}$ for $j \geq 1$. Hence the proof is complete. \square

5. HALF-REGION DEPTH OVER FINITE SUBSETS

In order to make half-region depth more amenable to discrete computations we now define half-region depth over finite sets, and prove a uniform consistency result in this setting.

As before we assume $X := \{X(t) = X_t : t \in T\}$ is a stochastic process on the probability space (Ω, \mathcal{F}, P) , all of whose sample paths are in $M(T)$.

If $h \in M(T)$ we define the half-region P -depth of h with respect to $J \subseteq T$ to be

$$(77) \quad D_J(h) = \min\{P(X \succeq_J h), P(X \preceq_J h)\},$$

where $h_1 \succeq_J h_2$ ($h_1 \preceq_J h_2$) holds for functions h_1, h_2 defined on T if $h_1(t) \geq h_2(t)$ ($h_1(t) \leq h_2(t)$) for all $t \in J$.

Let X_1, X_2, \dots be i.i.d. copies of the process X , and assume X, X_1, X_2, \dots are defined on (Ω, \mathcal{F}, P) suitably enlarged, if necessary, and that all sample paths of each X_j are in $M(T)$. Then, the empirical half-region depth of $h \in M(T)$ over a set $J \subseteq T$ is given by

$$(78) \quad D_{n,J}(h) = \min\left\{\frac{1}{n} \sum_{j=1}^n I(X_j \succeq_J h), \frac{1}{n} \sum_{j=1}^n I(X_j \preceq_J h)\right\},$$

For $h \in M(T)$ and J any finite subset of T , the probabilities in (77) are defined, and the events in (78) are in \mathcal{F} . Therefore, the classical law of large numbers implies with probability one

$$(79) \quad \lim_{n \rightarrow \infty} |D_{n,J}(h) - D_J(h)| = 0.$$

The next theorem refines (79) to be uniform over h and J , as long as $J \in \mathcal{J}_r$, where for each integer $r \geq 1$,

$$\mathcal{J}_r = \{J \subseteq T : \#J \leq r\},$$

and $\#J$ denotes the cardinality of the set J .

Theorem 3. Let X, X_1, X_2, \dots be i.i.d. copies of the stochastic process $X = \{X(t) : t \in T\}$ defined on the probability space (Ω, \mathcal{F}, P) , and all of whose sample paths are in the linear space $M(T)$. Let

$$(80) \quad \mathcal{C} = \{C_{t,y} : t \in T, y \in \mathbb{R}\},$$

where $C_{t,y} = \{z \in M(T) : z(t) \leq y\}$, and assume the empirical CLT holds with respect to the probability $\mathcal{L}(X)$ over \mathcal{C} . Then, for every integer $r \geq 1$ fixed we have with probability one that

$$(81) \quad \lim_{n \rightarrow \infty} \left[\sup_{h \in M(T)} \sup_{J \in \mathcal{J}_r} |D_{n,J}(h) - D_J(h)| \right]^* = 0.$$

Remark 6. The implication in (81) does not follow from the corresponding finite dimensional result for half-region depth since the finite set J is not fixed, but it is also the case that the assumption of an empirical CLT over \mathcal{C} is non-trivial. Fortunately [KKZ13] and [KZ13b] provide many examples of processes that satisfy this empirical CLT, and to which Theorem 3 applies. These include a broad collection of Gaussian processes, compound Poisson processes, stationary independent increment stable processes, and martingales. Moreover, if $J : \Theta \rightarrow \mathcal{J}_r$, then

$$\left[\sup_{h \in M(T)} \sup_{\theta \in \Theta} |D_{n,J(\theta)}(h) - D_{J(\theta)}(h)| \right]^* \leq \left[\sup_{h \in M(T)} \sup_{J \in \mathcal{J}_r} |D_{n,J}(h) - D_J(h)| \right]^*,$$

and hence it is immediate that (81) holds when the choice of J is arbitrarily parameterized by Θ as long as $\#J(\theta) \leq r, \theta \in \Theta$.

Proof. Since (24) holds, we have

$$|D_{n,J}(h) - D_J(h)| \leq A_n(J, h) + B_n(J, h),$$

where

$$A_n(J, h) = \left| \frac{1}{n} \sum_{j=1}^n I(X_j \succeq_J h) - P(X \succeq_J h) \right|$$

and

$$B_n(J, h) = \left| \frac{1}{n} \sum_{j=1}^n I(X_j \preceq_J h) - P(X \preceq_J h) \right|.$$

Therefore, sub-additivity of measurable cover functions implies (81) will follow once we verify that with probability one

$$\lim_{n \rightarrow \infty} \left[\sup_{h \in M(T)} \sup_{J \in \mathcal{J}_r} A_n(J, h) \right]^* = \lim_{n \rightarrow \infty} \left[\sup_{h \in C[0,1]} \sup_{J \in \mathcal{J}_r} B_n(J, h) \right]^* = 0.$$

Fix an integer $r \geq 1$, and set

$$\phi(u_1, \dots, u_r) = \min\{u_1, \dots, u_r\}.$$

Then, for $J \in \mathcal{J}_r$, $h \in M(T)$, and $j = 1, \dots, r$, define

$$f_j = I_{C_{t_j, h(t_j)}},$$

which implies

$$D_{J,h} = \{z \in M(T) : z(t) \leq h(t), t \in J\} = \cap_{j=1}^r C_{t_j, h(t_j)},$$

and

$$I_{D_{J,h}} = \phi(f_1, \dots, f_r).$$

Since \mathcal{C} is a Donsker class with respect to P and $r \geq 1$ is fixed, Theorem 2.10.6 of [vdVW96] implies

$$\mathcal{D} = \{D_{J,h} : J \in \mathcal{J}_r, h \in M(T)\}$$

is also P -Donsker with respect to P . Thus by Lemma 2.10.14 of [vdVW96] we have almost surely that

$$(82) \quad \lim_{n \rightarrow \infty} \left[\sup_{h \in M(T)} \sup_{J \in \mathcal{J}_r} B_n(J, h) \right]^* = 0.$$

Since $\tilde{\mathcal{D}} = \{D_{J,h}^c : D_{J,h} \in \mathcal{J}_r\}$ is then also a Donsker class, the above argument implies that

$$(83) \quad \lim_{n \rightarrow \infty} \left[\sup_{h \in M(T)} \sup_{J \in \mathcal{J}_r} A_n(J, h) \right]^* = 0.$$

Combining (82) and (83) implies (81). □

REFERENCES

- [CC14] Anirvan Chakraborty and Probal Chaudhuri, *On data depth in infinite dimensional spaces*, Annals of the Institute of Statistical Mathematics (2014).
- [DGC11] Subhajit Dutta, Anil K. Ghosh, and Probal Chaudhuri, *Some intriguing properties of Tukey's half-space depth*, Bernoulli **17** (2011), no. 4, 1420–1434. MR 2854779 (2012k:62147)
- [Dud99] R. M. Dudley, *Uniform central limit theorems*, Cambridge Studies in Advanced Mathematics, vol. 63, Cambridge University Press, Cambridge, 1999. MR 1720712 (2000k:60050)
- [HJ74] Jørgen Hoffmann-Jørgensen, *Sums of independent Banach space valued random variables*, Studia Math. **52** (1974), 159–186. MR 50 #8626
- [JM75] Naresh C. Jain and Michael B. Marcus, *Integrability of infinite sums of independent vector-valued random variables*, Trans. Amer. Math. Soc. **212** (1975), 1–36. MR 0385995 (52 #6854)
- [KKZ13] James Kuelbs, Thomas Kurtz, and Joel Zinn, *A CLT for empirical processes involving time-dependent data*, Ann. Probab. **41** (2013), no. 2, 785–816. MR 3077526
- [Kue78] J. Kuelbs, *Some exponential moments of sums of independent random variables*, Trans. Amer. Math. Soc. **240** (1978), 145–162. MR 0517296 (58 #24440)
- [KZ13a] James Kuelbs and Joel Zinn, *Concerns with functional depth*, ALEA Lat. Am. J. Probab. Math. Stat. **10** (2013), no. 2, 831–855.
- [KZ13b] ———, *Empirical quantile clts for time-dependent data*, High Dimensional Probability VI, The Banff Volume (Jan Rosinski Jon Wellner Christian Houdre, David Mason, ed.), Progress in Probability, Birkhauser, vol. 66, Springer Basel, 2013, pp. 167–194 (English).
- [LPR11] Sara López-Pintado and Juan Romo, *A half-region depth for functional data*, Comput. Statist. Data Anal. **55** (2011), no. 4, 1679–1695.
- [LT91] Michel Ledoux and Michel Talagrand, *Probability in Banach spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 23, Springer-Verlag, Berlin, 1991, Isoperimetry and processes. MR 1102015 (93c:60001)
- [Oss87] Mina Ossander, *A central limit theorem under metric entropy with L_2 bracketing*, Ann. Probab. **15** (1987), no. 3, 897–919. MR 893905 (88k:60067)
- [vdVW96] Aad W. van der Vaart and Jon A. Wellner, *Weak convergence and empirical processes*, Springer Series in Statistics, Springer-Verlag, New York, 1996, With applications to statistics. MR 1385671 (97g:60035)

JAMES KUELBS, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN, MADISON, WI 53706-1388

E-mail address: kuelbs@math.wisc.edu

JOEL ZINN, DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843-3368

E-mail address: jzinn@math.tamu.edu